

**IN THE CLAIMS:**

Please find below a listing of all pending claims. The statuses of the claims are set forth in parentheses. For those currently amended claims, underlined emphasis indicates insertions and ~~strikethrough~~ emphasis (and/or double brackets) indicates deletions.

1. (previously canceled).
2. (canceled).
3. (previously presented) A text change frame detection apparatus that selects a plurality of video frames including text contents from given video frames, said apparatus comprising:
  - a first frame removing unit to remove redundant video frames from the given video frames;
  - a second frame removing unit to remove video frames that do not contain a text area from the given video frames;
  - a third frame removing unit to detect and remove redundant video frames caused by image shifting from the given video frames; and
  - an output unit to output remaining video frames as candidate text change frames,wherein the second frame removing unit includes:
  - a fast and simple image binarization unit to generate a first binary image of a video frame of the given video frames;
  - a text line region determination unit to determine a position of a text line region by using a horizontal projection and a vertical projection of the first binary image;
  - a rebinarization unit to generate a second binary image of every text line region;

a text line confirmation unit to determine validity of a text line region by using a difference between the first binary image and the second binary image and a fill rate of a number of foreground pixels in the text line region to a total number of pixels in the text line region; and

a text frame verification unit to confirm whether a set of continuous video frames are non-text frames that do not contain a text area by using a number of valid text line regions in the set of continuous video frames.

4. (previously presented) A text change frame detection apparatus that selects a plurality of video frames including text contents from given video frames, said apparatus comprising:

a first frame removing unit to remove redundant video frames from the given video frames;

a second frame removing unit to remove video frames that do not contain a text area from the given video frames;

a third frame removing unit to detect and remove redundant video frames caused by image shifting from the given video frames; and

an output unit to output remaining video frames as candidate text change frames,

wherein the third frame removing unit includes:

a fast and simple image binarization unit to generate binary images of two video frames of the given video frames;

a text line vertical position determination unit to determine a vertical position of every text line region by using horizontal projections of the binary images of the two video frames;

a vertical shifting detection unit to determine a vertical offset of image shifting between the two video frames and a similarity of the two video frames in a vertical direction by using correlation between the horizontal projections; and

a horizontal shifting detection unit to determine a horizontal offset of the image shifting and a similarity of the two video frames in a horizontal direction by using correlation between vertical projections of every text line in the binary images of the two video frames,

and the third frame removing unit to remove a similar video frame as a redundant video frame caused by the image shifting.

5. (previously presented) A text change frame detection apparatus that selects a plurality of video frames including text contents from given video frames, said apparatus comprising:

an image block validation unit to calculate a mean value and a variance of a gray level of each of two image blocks in the same position in two video frames of the given video frames, and to determine the two image blocks are a valid block pair that has an ability to show a change of image contents if at least one of two variances of the two image blocks is greater than a first threshold, or if the two variances are smaller than the first threshold and an absolute difference of two mean values of the two image blocks is greater than a second threshold;

an image block similarity measurement unit to calculate a similarity of two image blocks of the valid block pair and to determine whether the two image blocks are similar;

a frame similarity judgment unit to determine whether the two video frames are similar by using a ratio of a number of similar image blocks to a total number of valid block pairs; and

an output unit to output remaining video frames after a similar video frame is removed, as candidate text change frames.

6. (previously presented) A text change frame detection apparatus that selects a plurality of video frames including text contents from given video frames, said apparatus comprising:

a fast and simple image binarization unit to generate a first binary image of a video frame of the given video frames;

a text line region determination unit to determine a position of a text line region by using a horizontal projection and a vertical projection of the first binary image;

a rebinarization unit to generate a second binary image of every text line region;

a text line confirmation unit to determine validity of a text line region by using a difference between the first binary image and the second binary image and a fill rate of a number of foreground pixels in the text line region to a total number of pixels in the text line region;

a text frame verification unit to confirm whether a set of continuous video frames are non-text frames that do not contain a text area by using a number of valid text line regions in the set of continuous video frames; and

an output unit to output remaining video frames after the non-text frames are removed, as candidate text change frames.

7. (previously presented) A text change frame detection apparatus that selects a plurality of video frames including text contents from given video frames, said apparatus comprising:

a fast and simple image binarization unit to generate binary images of two video frames of the given video frames;

a text line vertical position determination unit to determine a vertical position of every text line region by using horizontal projections of the binary images of the two video frames;

a vertical shifting detection unit to determine a vertical offset of image shifting between the two video frames and a similarity of the two video frames in a vertical direction by using correlation between the horizontal projections;

a horizontal shifting detection unit to determine a horizontal offset of the image shifting and a similarity of the two video frames in a horizontal direction by using correlation between vertical projections of every text line in the binary images of the two video frames; and

an output unit to output remaining video frames after a similar video frame is removed, as candidate text change frames.

8 - 15. (canceled).

16. (previously canceled)

17. (canceled).

18. (previously presented) A computer-readable storage medium storing a program used to direct a computer, that selects a plurality of video frames including text contents from given video frames, to perform a process comprising:

removing redundant video frames from the given video frames;

removing video frames that do not contain a text area from the given video frames;

detecting and removing redundant video frames caused by image shifting from the given video frames; and

outputting remaining video frames as candidate text change frames, wherein the removing video frames that do not contain the text area includes:

generating a first binary image of a video frame of the given video frames;

determining a position of a text line region by using a horizontal projection and a vertical projection of the first binary image;

generating a second binary image of every text line region;

determining validity of a text line region by using a difference between the first binary image and the second binary image and a fill rate of a number of

foreground pixels in the text line region to a total number of pixels in the text line region; and

confirming whether a set of continuous video frames are non-text frames that do not contain a text area by using a number of valid text line regions in the set of continuous video frames.

19. (previously presented) A computer-readable storage medium storing a program used to direct a computer, that selects a plurality of video frames including text contents from given video frames, to perform a process comprising:

removing redundant video frames from the given video frames;

removing video frames that do not contain a text area from the given video frames;

detecting and removing redundant video frames caused by image shifting from the given video frames; and

outputting remaining video frames as candidate text change frames, wherein the detecting and removing redundant video frames caused by image shifting includes:

generating binary images of two video frames of the given video frames;

determining a vertical position of every text line region by using horizontal projections of the binary images of the two video frames;

determining a vertical offset of image shifting between the two video frames and a similarity of the two video frames in a vertical direction by using correlation between the horizontal projections; and

determining a horizontal offset of the image shifting and a similarity of the two video frames in a horizontal direction by using correlation between vertical projections of every text line in the binary images of the two video frames, and the detecting and removing redundant video frames removes a similar video frame as a redundant video frame caused by the image shifting.

20. (previously presented) A computer-readable storage medium storing a program used to direct a computer, that selects a plurality of video frames including text contents from given video frames, to perform a process comprising:

calculating a mean value and a variance of a gray level of each of two image blocks in the same position in two video frames of the given video frames, and determining the two image blocks are a valid block pair that has an ability to show a change of image contents if at least one of two variances of the two image blocks is greater than a first threshold, or if the two variances are smaller than the first threshold and an absolute difference of two mean values of the two image blocks is greater than a second threshold;

calculating a similarity of two image blocks of the valid block pair and determining whether the two image blocks are similar;

determining whether the two video frames are similar by using a ratio of a number of similar image blocks to a total number of valid block pairs; and

outputting remaining video frames after a similar video frame is removed, as candidate text change frames.

21. (original) A computer-readable storage medium storing a program used to direct a computer, that selects a plurality of video frames including text contents from given video frames, to perform a process comprising:

generating a first binary image of a video frame of the given video frames;

determining a position of a text line region by using a horizontal projection and a vertical projection of the first binary image;

generating a second binary image of every text line region;

determining validity of a text line region by using a difference between the first binary image and the second binary image and a fill rate of a number of foreground pixels in the text line region to a total number of pixels in the text line region;

confirming whether a set of continuous video frames are non-text frames that do not contain a text area by using a number of valid text line regions in the set of continuous video frames; and

outputting remaining video frames after the non-text frames are removed, as candidate text change frames.

22. (original) A computer-readable storage medium storing a program used to direct a computer, that selects a plurality of video frames including text contents from given video frames, to perform a process comprising:

generating binary images of two video frames of the given video frames;

determining a vertical position of every text line region by using horizontal projections of the binary images of the two video frames;

determining a vertical offset of image shifting between the two video frames and a similarity of the two video frames in a vertical direction by using correlation between the horizontal projections;

determining a horizontal offset of the image shifting and a similarity of the two video frames in a horizontal direction by using correlation between vertical projections of every text line in the binary images of the two video frames; and

outputting remaining video frames after a similar video frame is removed, as candidate text change frames.

23 - 30. (canceled).

31. (previously canceled)

32. (canceled).